



UNIVERSIDAD
COMPLUTENSE
MADRID

Proyecto de Innovación y Mejora de la Calidad Docente

Convocatoria 2018/19

Nº 20

Lo que no conocemos de nuestros estudiantes de CC de la Salud y nos revelan los rastros digitales. *Learning Analytics*: herramienta con potencial transformador para mejorar nuestra docencia.

M^a Pilar Álvarez Vázquez

Facultad de Medicina

Departamento de Biología Celular

1. OBJETIVOS PROPUESTOS EN LA PRESENTACIÓN DEL PROYECTO

El proyecto que presentamos tenía como objetivo principal:

- Proporcionar información a los docentes para la implementación de su metodología docente de acuerdo al análisis de los registros del Campus Virtual (CV) de la UCM y posterior interpretación sobre el comportamiento de sus estudiantes.

Este objetivo se estructuraba en los siguientes objetivos específicos:

1. Realizar un análisis descriptivo de los usos que hacen los estudiantes de varias asignaturas virtualizadas en cuatro grados de CC de la Salud: Medicina, Nutrición, Fisioterapia y Podología.
2. Establecer, si fuera posible, distintas tipologías de estudiantes de acuerdo a su comportamiento y uso de las asignaturas virtualizadas analizadas.
3. Correlacionar los posibles clusters de estudiantes encontrados en dicha asignatura virtualizada con otros parámetros, como su rendimiento académico.
4. Proporcionar al docente información sobre el dinamismo y la salud de su asignatura virtualizada.

Estos objetivos giraban en torno a las siguientes necesidades y problemas:

- Propuesta de valor 1: Ante el desconocimiento generalizado de los profesores sobre la actividad de sus estudiantes en Moodle, este proyecto busca proporcionar información más profunda al docente al respecto. Los profesores universitarios emplean plataformas digitales, Moodle en el caso de la UCM, para virtualizar asignaturas y espacios. Sin embargo, en su inmensa mayoría, desconocen la actividad que desarrollan sus estudiantes en el CV, tanto durante su periodo de impartición como después de haber terminado. Por ello nos parece esencial llevar a cabo un trabajo de minería o explotación de datos para tratar de proporcionar una lectura comprensible.
- Propuesta de valor 2: Emplear el CV no sólo como medio de enseñanza-aprendizaje sino también como fuente de información para implementar mejoras en la calidad de la enseñanza y del aprendizaje, en los siguientes cursos, en la docencia de la asignatura analizada, al permitir extraer conclusiones valiosas en base a datos objetivos y medibles. Los profesores universitarios dedican, curso tras curso, no poco tiempo y esfuerzo a los recursos que alojan en los espacios digitales y a diseñar tareas y actividades para el aprendizaje de sus estudiantes. Sin embargo, implementan su metodología docente al margen de datos reales de uso y hábitos de sus estudiantes en las asignaturas virtualizadas al tiempo que les impide conocer posibles puntos débiles y fortalezas en su asignaturas virtualizadas. En el actual PIMCD se ha procedido al análisis descriptivo de los hábitos de uso de los estudiantes de Organografía Microscópica Humana (OMH) y en la actualidad se están realizando análisis para ver si es posible determinar la existencia de clusters o categorías de alumnos de acuerdo a dichos hábitos.

En el nuevo PIMCD estaríamos en condiciones de ampliar potencialmente el estudio a nuevas asignaturas:

- del Grado en Medicina: Biología Celular, Embriología General e Histología Humana de primer curso;
- del Grado en Nutrición: Biología, de primer curso;
- del Grado en Fisioterapia: Anatomía III, de segundo curso;
- del Grado en Podología: Biología, de primer curso; Patología Médica II y

Biomecánica, de segundo curso y Salud Laboral de cuarto curso.

- Propuesta de valor 3: Desarrollar una metodología externa apropiada para analizar los informes que proporciona Moodle y otras fuentes de datos externas, y evaluar el comportamiento de los estudiantes de OMH.
- Propuesta de valor 4: Transferencia de recursos. El proyecto se plantea como un estudio progresivo en la selección, análisis y construcción de modelos a partir de los datos extraídos del CV. Es importante que una vez definidos los indicadores cuantitativos y cualitativos se elabore una guía sobre el desarrollo de la minería de datos llevada a cabo, de manera que sea extrapolable a otros espacios virtualizados, y que por tanto pueda ser empleada por otros profesores universitarios.

2. OBJETIVOS ALCANZADOS

Se han analizado todas de las asignaturas propuestas excepto la del Grado en Nutrición. En la Tabla 1 se indican los espacios del CV analizados.

Tabla 1. Espacios virtuales analizados.

Nombre del espacio	identificador	Curso	Grado en
Biología celular, Embriología general e Histología humana	17-80266	2017/18	Medicina
PRÁCTICAS Organografía Grupo IA 17/18	17-190652	2017/18	Medicina
PRÁCTICAS ORGANOGRAFÍA 2A. 2017-18	seminario-invest-3262-3	2017/18	Medicina
Anatomía humana III	17-141227	2017/18	Fisioterapia
Biomecánica	17-171904	2017/18	Podología
Salud laboral	17-473738	2017/18	Podología
Salud laboral	18-473738	2018/19	Podología
Organografía 18/19 grupo IA	18-93062	2018/19	Medicina
Prácticas Organografía 18/19 Grupo IA	seminario-invest-5914-19	2018/19	Medicina

Se han alcanzado todos los objetivos específicos propuestos. El trabajo en equipo interdisciplinar ha permitido progresar en el desarrollo del proyecto, mejorar aspectos que en el pimcd2017-31 fueron limitantes y establecer mejoras en el manual técnico elaborado en el mismo.

Como inconveniente a señalar, está el hecho de que las asignaturas anuales, como es el caso de OMH, tienen un calendario académico que dificulta su análisis dentro de los plazos establecidos para los pimcd. Así, su examen final de junio fue el 3/06, por lo que el análisis de la asignatura se solapa con la elaboración de la memoria, y obliga a dejar fuera la convocatoria de julio.

3. METODOLOGÍA EMPLEADA EN EL PROYECTO

- a) Encuestas: 1, al PDI sobre las asignaturas a analizar para conocer sus características, similitudes y diferencias en organización, metodologías y estructura de los espacios virtuales, empleando *Google Forms*; y 2, a los estudiantes sobre CV, acceso a herramientas web 2.0 y bibliotecas, conocimiento de TIC y características socio-demográficas. En hoja separada se recogió la relación de

encuestados y su identificador. Esto permitió unir la información de la encuesta 2 con los registros (*logs*) de CV y los datos de rendimiento académico.

- b) Fuentes de datos de los estudiantes analizados: encuesta 2, *logs* descargados de Moodle y rendimiento académico.
- c) Extracción de los *logs* de cada espacio virtual a analizar, descarga de calificaciones y obtención de las rutas de cada recurso y actividad. Las variables que se consideraron fueron el número de: accesos a CV, accesos a recursos (documentos en word, pdf, excel...) y a URL externas, autoevaluaciones y tareas realizadas, consultas a foros, creaciones/modificaciones en wikis, glosarios y foros.
- d) Estandarización de archivos, depuración y tratamiento anonimizado de los *logs*.
- e) Análisis con R. Creación de un aplicativo que permite cruzar resultados y gráficos con un número ilimitado de asignaturas y hasta 30 calificaciones en cada una, sistematizando las opciones de tratamiento de datos.
- f) Elaboración de una Declaración de Tipo de Datos (DTD) que permite documentar la estructura del documento XML.
- g) Análisis estadístico encuesta 2. Variables cualitativas mediante frecuencias, porcentajes y tablas de contingencia para visualizar la relación entre pares de variables, posteriormente analizada por estadístico exacto de Fisher. Se estudió la relación de las variables de la encuesta con los grados, comparando además un mismo curso de dos grados diferentes y dos cursos diferentes de un mismo grado.
- h) Análisis estadístico de uso de CV comparado con rendimiento académico. Estudio de cada asignatura. Correlaciones no paramétricas de Spearman entre calificaciones y accesos a CV, recursos y actividades. Se obtuvieron árboles de decisión con dos criterios de corte (minimización de la significatividad mediante el estadístico de Fisher-Snedecor y minimización de la varianza) para relacionar rendimiento académico con segmentos de estudiantes con una actividad similar en CV.
- i) Análisis estadístico de datos de las tres fuentes: Estudio de cada asignatura, relacionando las calificaciones de los estudiantes con sus características socio-demográficas, la consulta de bibliografía y usos de la biblioteca. Se relacionaron accesos a CV, recursos y actividades con la utilidad de los mismos en el aprendizaje manifestada en la encuesta B. Ello se hizo mediante los estadísticos de t-student y análisis de la varianza, en caso de normalidad y mediante contraste de distribuciones no paramétrico de Mann-Whitney en otro caso, aplicando Shapiro-Wilk.
- j) Visualización de datos: Curvas de frecuencias de acceso a CV, recursos y actividades en el curso académico; Diagramas de barras según acceso a CV, recursos y actividades día/semana; Mapas de calor según acceso a CV, recursos y actividades según día y hora; Diagramas de barras mostrando la distribución de las variables para visualizar diferencias entre grupos; Grafos para representar árboles de decisión; y Gráficos de estrella multivariados de Kiwi para representar accesos promedio a CV, recursos y actividades según segmentos de calificaciones.
- k) Fuentes de datos para comparar: registros generales de uso CV y registros de participación anonimizada de usuarios para los cursos oficiales de los grados en Medicina, Fisioterapia y Podología dados de alta en CV en los cursos 2017/18 y 2018/19, previa solicitud al VR. Obtención del número de alumnos en cada grado para dichos cursos

4. RECURSOS HUMANOS

En este pimcd han participado PDI, PAS y una estudiante.

- I. Bedmar Gómez ha preparado la encuesta 1 online sobre las características de las asignaturas a analizar, señalando las variables más interesantes para mostrar diferencias y semejanzas. PDI ha rellenado la encuesta. I. Bedmar ha elaborado los resultados.
- A. Álvarez Méndez, M.P. Álvarez Vázquez y J. Cristóbal Barrios han realizado búsquedas bibliográficas y discutido los trabajos de otros colegas.
- PDI ha elaborado la encuesta 2, de satisfacción de alumnos sobre el CV. La encuesta junto con los nombres de los alumnos encuestados pasó directamente a manos de M.C. Bravo Llatas, quien la procesó y realizó los análisis indicados en el apartado anterior.
- PDI, cada profesor ha proporcionado sus listas de clase, las fechas importantes en la materia, así como el rendimiento académico de sus estudiantes.
- M. Castañer de Diego y M. Parra Bronchalo han recopilado las rutas de cada fichero, actividad y herramientas empleadas en las asignaturas y en los seminarios.
- De cara a tener mejores referencias con las que comparar los resultados obtenidos en cada asignatura, J. Merino Granizo ha recopilado los siguientes registros almacenados en Moodle, previa solicitud y autorización del Vicerrector de Tecnologías de la Información (VRTI):
 - Obtención de la información de participación anonimizada de los usuarios para los cursos oficiales de los grados de Medicina, Podología y Fisioterapia de los cursos 2017/2018 y 2018/2019 hasta el día 2 de mayo y que estuvieran dados de alta en *Moodle*.
 - Obtención del número de alumnos de cada grado de los cursos 2017/2018 y 2018/2019 hasta el día 2 de mayo.
- A partir del manual técnico elaborado en el pimcd2017-31, J. Cristóbal Barrios ha depurado y analizado de forma anonimizada todos los *logs* de *Moodle* de las asignaturas y espacios analizados, los listados de estudiantes y de calificaciones, obteniendo distintos tipos de gráficas.
- M.C. Bravo Llatas ha analizado estadísticamente los datos agregados de cada asignatura, realizado igualmente distintos tipos de gráficas, correlaciones y estudios comparados intra- e intertitulación.
- El conjunto del PDI y PAS se ha encargado de debatir las características de la materia, de planificar las etapas de la investigación, de discutir la categorización de los registros, y de analizar resultados y extraer conclusiones.
- Estudiante, PDI y PAS han contribuido a la redacción de la presente memoria.

5. DESARROLLO DE LAS ACTIVIDADES

En un primer momento se abordó el estudio de las características de las asignaturas a analizar desde el punto de vista de su organización académica y de la estructura y

desarrollo de los espacios virtualizados. Partiendo del análisis del 2017/18, se introdujeron cambios en la planificación de Salud laboral del 2018/19.

El estudio de cada asignatura y espacio virtualizado se ha llevado a cabo de forma progresiva. En una primera fase se estudiaron los espacios del curso 2017/18 ya que los datos estaban completos. Seguidamente se sumó el espacio de Salud laboral del curso actual, para terminar con los dos espacios de OMH, uno de prácticas y otro de teoría.

A diferencia del pimcd2017-31 en donde los registros se obtuvieron directamente por J. Merino del área de CV, en el actual los *logs* se han descargado directamente de *Moodle*. El procedimiento es sencillo y cómodo (Figura 1). Además tiene la ventaja de que este tipo de registros almacenan información valiosa que no se obtiene por el procedimiento seguido en el pimcd2017-31¹.



Figura 1. Procedimiento para descargar los registros en Moodle.

El análisis de las asignaturas se ha hecho partiendo del manual técnico² elaborado por J. Cristóbal Barrios para la asignatura OMH, adaptándolo a los nuevos espacios cuyas características son diferentes.

El volumen de registros descargados de los espacios analizados y manejados por J. Cristóbal ha sido de 183.000 logs de 2017/18 y 218.000 logs en 2018/19.

La visualización de los datos puso de manifiesto la conveniencia de poder comparar los resultados de las asignaturas virtualizadas con una muestra complutense más limitada que la que proporciona *Google Analytics* (todos los espacios UCM en *Moodle*), por lo que se solicitó al VRTI información sobre las titulaciones de las asignaturas analizadas, con lo que podríamos tener una visión más adecuada. Los registros que J. Merino proporcionó desde eCampus están actualmente pendientes de

¹ Álvarez M.P. *Rastreando Moodle. Logs: limitaciones y dificultades*.

<https://eprints.ucm.es/51122/>

² Álvarez M.P. et al. 2018. *Learning Analytics: Campus Virtual como fuente de información para conocer a nuestros estudiantes y mejorar la calidad docente*. <https://eprints.ucm.es/48136/>

ser procesados por lo que las gráficas comparadas que se incluyen en la presente memoria siguen haciendo referencia a UCM global y no a una titulación concreta, como esperamos poder hacer en breve. El total de registros asciende más de 10.000.000 de los cuales 993.000 corresponden a Fisioterapia, 983.000 a Podología y 8.072.000 a Medicina.

A continuación se expone una parte de los hallazgos más significativos. En relación a la actividad desarrollada por los estudiantes en los espacios virtuales cabe reseñar que en todos los casos está estrechamente ligada a las fechas de pruebas, cuestionarios, entrega de tareas y exámenes. Esto se constata tanto en el número de visitas a los espacios, como de accesos a recursos, a foros y a URL externas (Figura 2).

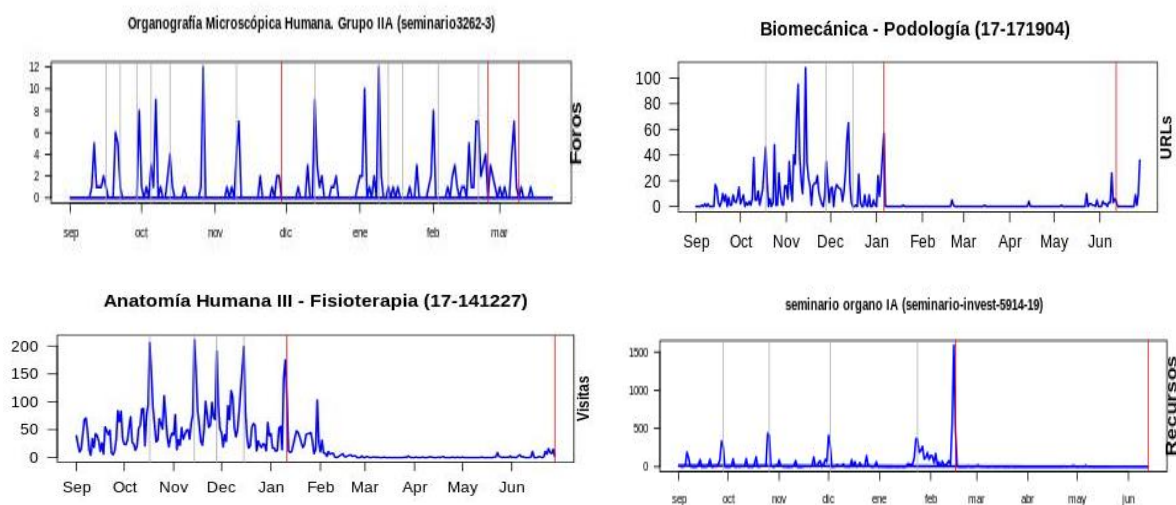


Figura 2. Acceso medio al CV, recursos, foros y URLs en cuatro espacios. Las líneas rojas señalan fechas de exámenes y las grises de cuestionarios, pruebas o tareas.

Los cambios en la planificación docente de una asignatura se reflejan claramente en los registros en Moodle. Así, por ejemplo, en 2018/19 se ofertaron en Salud laboral nuevas actividades (cuestionarios, wiki, glosario) que condicionaron el comportamiento de los estudiantes en el CV (Figura 3).

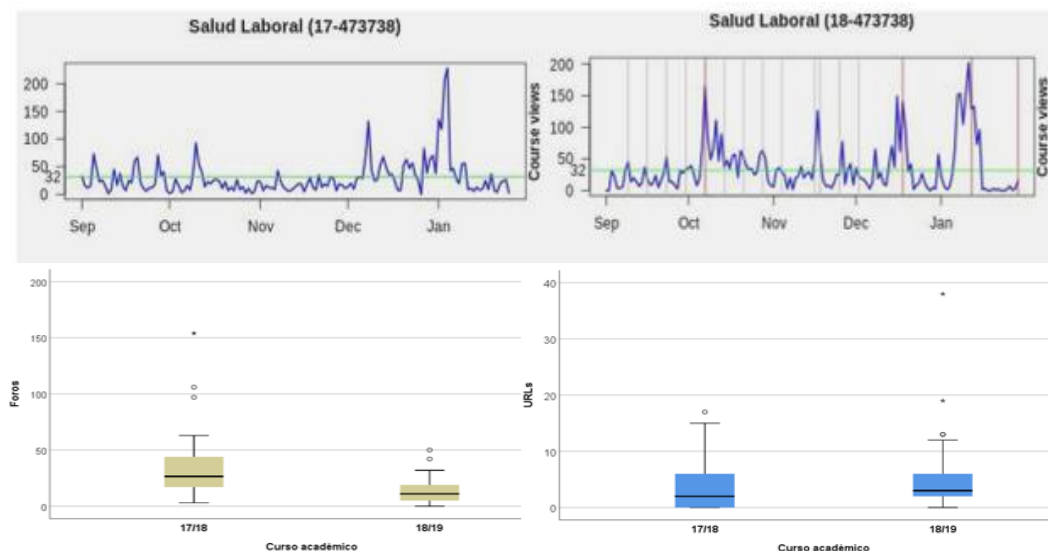


Figura 3. Actividad comparada en Salud laboral en 2017/18 y 2018/19.

Los estudiantes de cada grupo muestran comportamientos heterogéneos en el CV y ello repercute en su rendimiento académico, como se pone de manifiesto en buena parte de las calificaciones obtenidas en las distintas asignaturas. A modo de ejemplo se recogen los árboles de decisión para las calificaciones de Febrero en Anatomía III en los que se constata que los alumnos que hacen un mayor uso de los recursos que el docente facilita y de las actividades propuestas obtienen mejores notas (Figura 4).

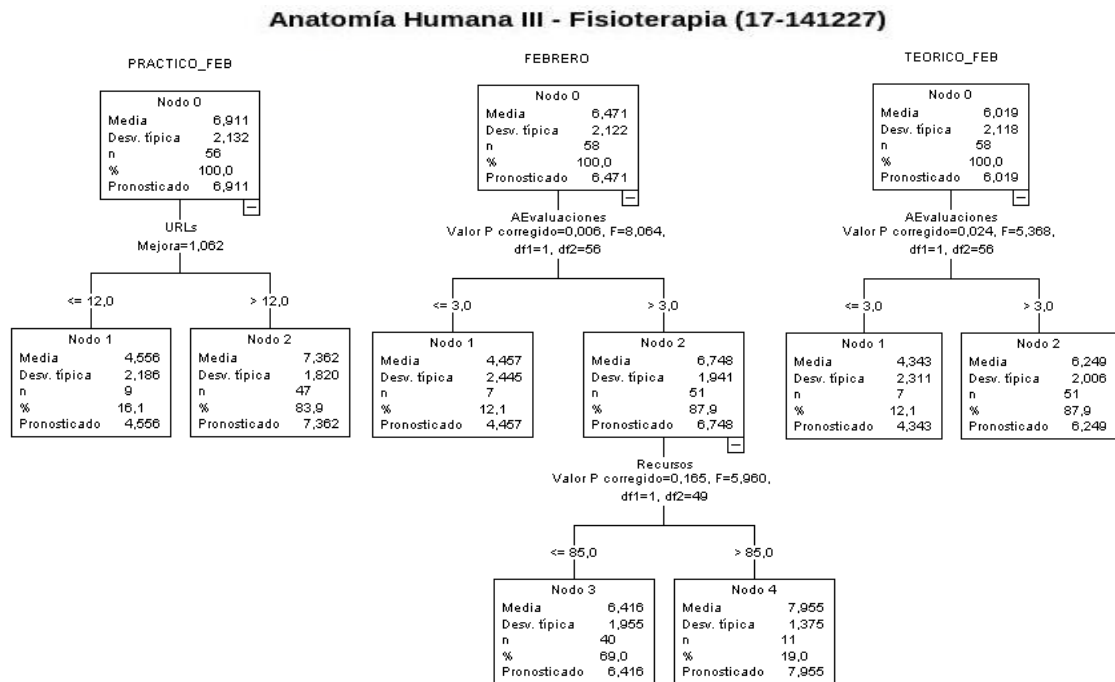


Figura 4. Los árboles muestran que los estudiantes de Anatomía III con un uso mayor del CV obtienen mejores calificaciones (criterio de la F de Fisher-Snedecor y estimaciones de calidad (varianza ponderada en los nodos finales) de 3,403; 3,514 y 4,023, respectivamente).

La existencia de *clústeres* también se puede visualizar en gráficas en estrella, como se observa en la Figura 5 que compara la nota final de prácticas en dos cursos sucesivos.

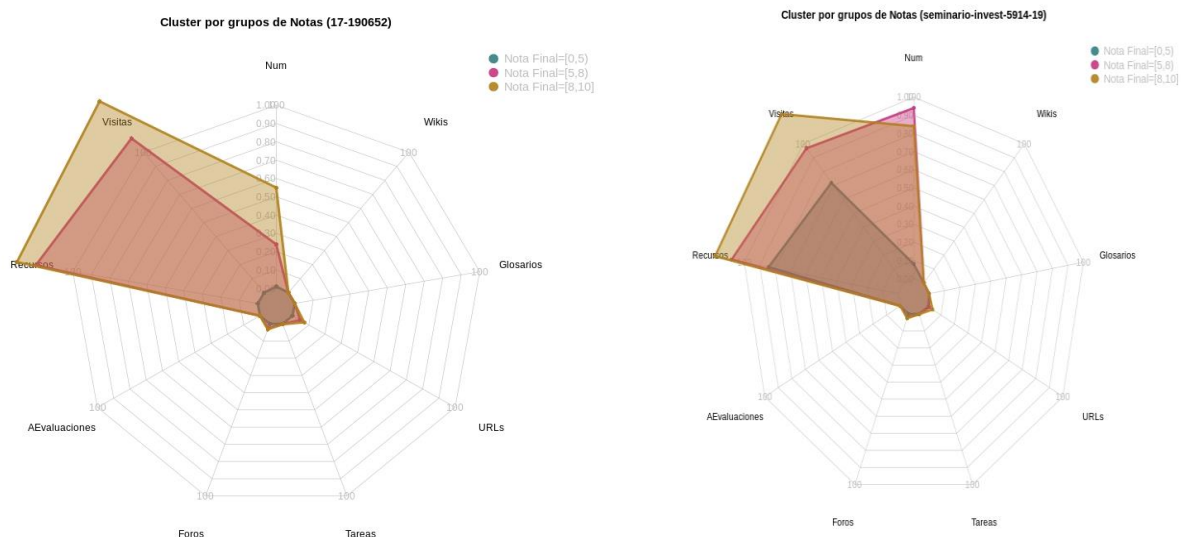


Figura 5. Comparativa de clústeres en el mismo espacio de prácticas de OMH en 2017/18 y 2018/19.

Los resultados obtenidos al cruzar la encuesta 2, el uso del CV y el rendimiento académico mediante contraste de distribuciones de Mann-Whitney revelan que los encuestados de Biomecánica que trabajan obtienen peores notas y, en el caso de los de OMH, los que emplean préstamo bibliotecario consiguen mejores notas en la evaluación continua en prácticas (Figura 6).

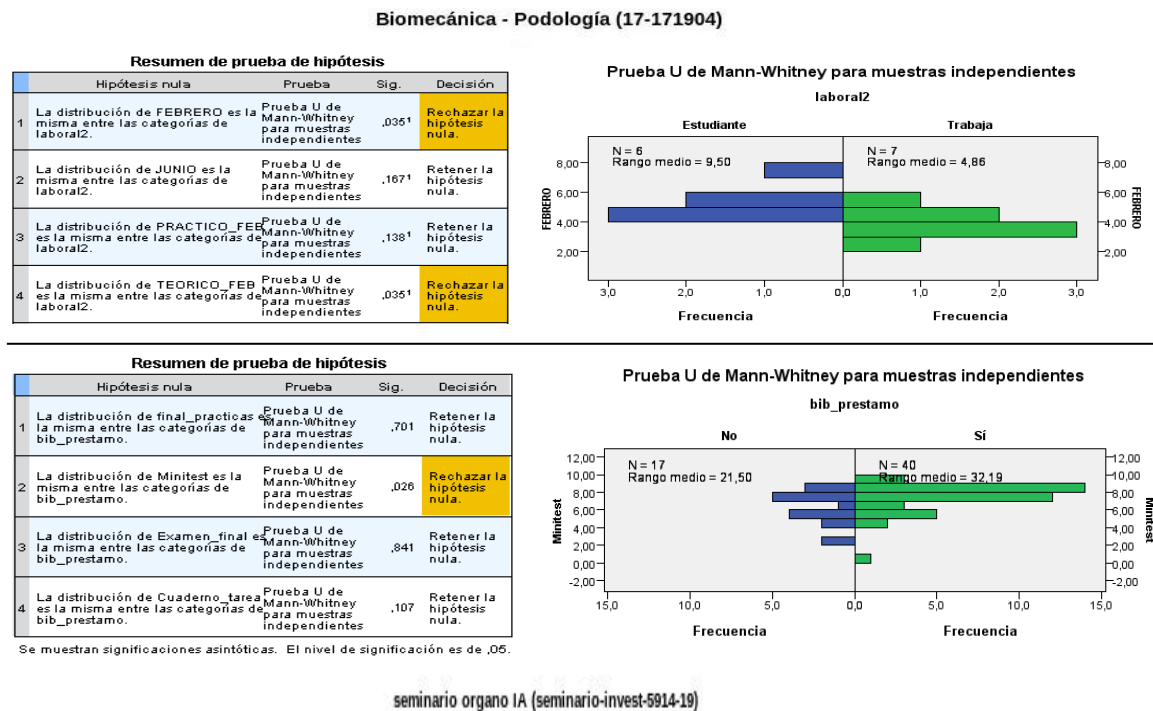


Figura 6. Contraste de distribuciones para encuestados de Biomecánica y de OMH.

El estudio comparado entre encuestados de 2º y 4º curso de Podología muestra que el uso de bases científicas es mayor en los de 4º curso (25,6% en 4º vs 0% en 2º, $p=0,04$), lo que podría explicarse por el trabajo fin de grado. Por su parte, el estudio comparado entre encuestados de 2º curso de Medicina y de Podología revela diferencias significativas en distintos ítems, entre ellos, la mejor valoración que hacen los primeros de las tareas grupales para su aprendizaje y de los cuestionarios en el CV frente a los segundos (97,2% vs 64,3%, $p=0,001$; 82,5% vs 57,1%, $p=0,05$) y o el mayor uso de redes sociales de los futuros podólogos (50,0% vs 5,3%, $p<0,001$).

ANEXOS

MANUAL TÉCNICO DE ANÁLISIS DE CURSOS DE MOODLE SOBRE R

Este documento sirve para mostrar la forma de proceder por cualquier usuario para realizar la evaluación de los *logs* proporcionados por el CV de la UCM (*Moodle*), dentro de la herramienta desarrollada. Esta herramienta hace uso de un script escrito en el lenguaje de programación R, capaz de correr en cualquier tipo de ordenador (Windows, Linux, Mac), que permite recuperar información de distintos documentos y generar gráficos en archivos individualizados en formato jpg así como realizar tratamientos estadísticos complejos. Los distintos archivos que procesar y demás datos de entrada se han “empaquetado” en un documento en formato xml para facilitar la información necesaria para ser procesada. Este archivo puede ser modificado con cualquier editor, aunque sería recomendable utilizar un programa para generarlo de forma automatizada. Este archivo está caracterizado siguiendo una estructura denotada en un archivo DTD (Document Type Definition³):

```
<!DOCTYPE cursos [  
<!ELEMENT cursos (curso+)>  
<!ELEMENT curso (id,anoaca,author?,title?,url?,logfile,logvers,studentsfile,gradesfile,fec-  
ini,fec-fin,grades,parciales?,autoevaluaciones?,foros?,tareas?)>  
<!ELEMENT id (#PCDATA)>  
<!ELEMENT anoaca (#PCDATA)>  
<!ELEMENT author (#PCDATA)>  
<!ELEMENT title (#PCDATA)>  
<!ELEMENT url (#PCDATA)>  
<!ELEMENT logfile (#PCDATA)>  
<!ELEMENT logvers (#PCDATA)>  
<!ELEMENT studentsfile (#PCDATA)>  
<!ELEMENT gradesfile (#PCDATA)>  
<!ELEMENT fec-ini (#PCDATA)>  
<!ELEMENT fec-fin (#PCDATA)>  
<!ELEMENT grades (main_grade,main_grade2?,grade+)>  
<!ELEMENT main_grade (name,passgrade?)>  
<!ELEMENT main_grade2 (name,passgrade?)>  
<!ELEMENT grade (name,passgrade?)>  
<!ELEMENT parciales (parcial*)>  
<!ELEMENT autoevaluaciones (autoevaluacion*)>  
<!ELEMENT foros (foro*)>  
<!ELEMENT tareas (tarea*)>  
<!ELEMENT parcial (#PCDATA)>  
<!ELEMENT autoevaluacion (#PCDATA)>  
<!ELEMENT tarea (#PCDATA)>  
<!ELEMENT foro (#PCDATA)>  
<!ELEMENT name (#PCDATA)>  
<!ELEMENT passgrade (#PCDATA)>  
>]
```

El archivo XML (*eXtensible Markup Language*⁴) actúa como índice y contiene la información que precisa el aplicativo para gestionar los datos y elaborar los resultados y estadísticas:

³ Document Type Definition <https://www.w3.org/QA/2002/04/valid-dtd-list.html>

⁴ *eXtensible Markup Language*
https://developer.mozilla.org/es/docs/Web/XML/Introducción_a_XML

```

<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE address SYSTEM "index.dtd">
<cursos>
  <curso>
    <id>17-190652</id>
    <anoaca>2017/18</anoaca>
    <author>Pilar Álvarez</author>
    <title>Organografía Microscópica Humana. Grupo IA</title>
    <url></url>
    <logfile>./17-190652 Registros.csv</logfile>
    <logvers>2.9.l</logvers>
    <studentsfile>./17-190652 Alumnos.csv</studentsfile>
    <gradesfile>./17-190652 Calificaciones.csv</gradesfile>
    <resourcesfile></resourcesfile>
    <fec-ini>2017-09-07</fec-ini>
    <fec-fin>2018-06-29</fec-fin>
    <grades>
      <main_grade>
        <name>Junio</name>
        <passgrade>5</passgrade>
      </main_grade>
      <main_grade2>
        <name>Julio</name>
        <passgrade>5</passgrade>
      </main_grade2>
      <grade>
        <name>Minitests (Real)</name>
        <passgrade>5</passgrade>
      </grade>
      <grade>
        <name>Examen final (Real)</name>
        <passgrade>5</passgrade>
      </grade>
      <grade>
        <name>Nota del cuaderno (Real)</name>
        <passgrade>5</passgrade>
      </grade>
      <grade>
        <name>NOTA FINAL DE PRÁCTICAS (Real)</name>
        <passgrade>5</passgrade>
      </grade>
    </grades>
    <parciales>
      <parcial>2018-02-22</parcial>
      <parcial>2018-06-29</parcial>
      <parcial></parcial>
      <parcial></parcial>
      <parcial></parcial>
      <parcial></parcial>
    </parciales>
    <autoevaluaciones>
      <autoevaluacion>2017-09-20</autoevaluacion>
      <autoevaluacion>2017-09-26</autoevaluacion>
      <autoevaluacion>2018-01-23</autoevaluacion>
      <autoevaluacion>2018-02-12</autoevaluacion>
      <autoevaluacion>2018-03-13</autoevaluacion>
    </autoevaluaciones>
  </curso>
</cursos>

```

```
        <foros>
        </foros>
        <tareas>
            <tarea>2018-02-12</tarea>
            <tarea>2019-03-13</tarea>
        </tareas>
    </curso>
</cursos>
```

1. El curso o cursos que evaluar:<curso>

- a. <id>: Identificación dentro del CV
- b. <anoaca>: Año académico (20XX/XX). Se mostrará en la cabecera de los gráficos
- c. <title>: Descripción del Curso
- d. <url>: URL del curso en el CV
- e. <logfile>: Fichero con el log de actividad del *moodle*. Extraído desde la propia herramienta moodle
- f. <logvers>: Versión de logs proporcionados por *moodle*. (2.6 /2.9)
- g. <studentsfile>: Fichero con los datos de los alumnos desde GEA.
- h. <gradesfile>: Fichero de calificaciones extraído desde *moodle*.
- i. <resourcesfile>: Fichero con el mapa de los recursos disponibles dentro del curso.
- j. <fec-ini>: Fecha de inicio del curso
- k. <fec-fin>: Fecha de finalización del curso
- l. <grades>: Descripción de las calificaciones empleadas en la asignatura. El fichero de calificaciones <logfile> puede contener más calificaciones, pero solo las indicadas en este punto serán consideradas. Cada asignatura dispondrá de dos calificaciones finales (<maingradeZ y <maingrade2> y hasta 30 calificaciones parciales que contemplen distintas actividades (parciales, autoevaluaciones, participación en foros, wikis, blogs). Para cada una de estas calificaciones debe detallarse un nombre (que deberá coincidir con el indicado en el fichero <gradesfile>) y una escala de valoración.
- m. <parciales>: Fechas de realización de los exámenes parciales
- n. <autoevaluaciones>: Fechas de realización de ejercicios de autoevaluación (evaluable).
- o. <foros>: Fechas de entrega de las participaciones en foros (evaluable).
- p. <tareas>: Fechas de entrega de las tareas (evaluable).

2. El archivo <logfile> contiene toda la información registrada en los ordenadores con la actividad de los alumnos y profesores. Habitualmente esta información es utilizada para detectar errores o mal funcionamiento de los módulos que componen el campus virtual, pero en este caso nos interesa para trazar la actividad de los usuarios dentro de la asignatura. En estos archivos no solo figura la información de accesos a las herramientas, también la actividad que se realiza sobre las propias herramientas. Por ejemplo, nos puede indicar el número de veces que un alumno realiza una tarea, pero también si ha sido corregida, visualizada por otros alumnos y/o profesores, el número de archivos que ha adjuntado en esa tarea, etc. Además,

en estos archivos figura la información de accesos de cualquier usuario (Profesor y Estudiante) dentro del curso. Por este motivo, es necesario realizar un filtro para procesar la información contenida en estos archivos. El contenido de estos archivos es:

id	time	userid	ip	course	module	cmid	action	url	info	idUs	DNI	nombre	apellidos	idEspacio	rol
467237550	1440412845	68801	17683114157	58232	course		0 view	view.php?id=58232	58232	5196	50306370	MARIA DEL PILAR	ALVAREZ VAZQUEZ	15-93062	Profesor
467238606	1440413691	68801	17683114157	58232	course		0 view	view.php?id=58232	58232	5196	50306370	MARIA DEL PILAR	ALVAREZ VAZQUEZ	15-93062	Profesor
467258644	1440436696	68801	95124095	58232	course		0 view	view.php?id=58232	58232	5196	50306370	MARIA DEL PILAR	ALVAREZ VAZQUEZ	15-93062	Profesor
467258669	1440436712	68801	95124095	58232	resource	1811862	view	view.php?id=1811862	1282173	5196	50306370	MARIA DEL PILAR	ALVAREZ VAZQUEZ	15-93062	Profesor

La estructura cambia con la versión de *moodle* pero, básicamente, nos fijamos en el módulo utilizado, la acción, y el rol del usuario para determinar la actividad de los usuarios dentro del curso. También es posible obtener estos datos desde la propia herramienta de *Moodle* como informe y extraerla en formato CSV (Comma-Separated Value⁵)

- El archivo permite evaluar varias asignaturas simultáneamente. El número de Parciales, Autoevaluaciones, Foros y Tareas es ilimitado.
- Las fechas de inicio y fin no tienen por qué coincidir con las de inicio y fin del curso. Pueden concretarse a un periodo y estudiarlo en profundidad.
- El campo <logvers> puede contener dos valores distintos: 2.6, 2.9 puesto que ambas versiones tienen estructuras distintas. También puede ser 2.9.I si el fichero log se extrae desde la herramienta de Moodle mediante informes.
- El archivo <studentsfile> debe contener los siguientes campos:
 - Nº de Matrícula
 - Nº de Convocatoria
 - Sexo
 - Apellidos
 - Nombre
 - DNI/NIF/Pasaporte
- El archivo <gradesfile> contiene las calificaciones obtenidas en cada uno de los parciales, así como las notas finales de las dos convocatorias febrero o junio y julio. Cada nota debe ir identificada con un título que deberá coincidir con las notas marcadas dentro de la estructura de <grades>
- El archivo <resourcefile> indica información sobre los recursos, su identidad dentro de Moodle, el formato del recurso (HTML, imagen, PDF, vídeo....)

Una vez preparados los archivos, la ejecución de la aplicación obtendrá distintos gráficos en formato jpg en archivos separados. Los gráficos que se obtienen son los siguientes:

⁵ Comma-Separated Value <https://tools.ietf.org/html/rfc4180>

A. **Utilización general de las herramientas en cada curso** (Tabla 1). Describe, de forma global, qué tanto por ciento de estudiantes acceden al curso y a cada una de las actividades.

Tabla 1. Uso del CV. Ejemplo: Salud laboral, cursos 2017/18 y 18/19.

	id	title	Visitas	Recursos	AEvaluaciones	Foros	Tareas	URLs	Glosarios	Wikis
1	17-473738	Salud Laboral	100	100	0.00	100.0	3.17	71.43	0.00	0.00
2	18-473738	Salud Laboral	100	100	96.55	93.1	91.38	87.93	87.93	48.28

B. **Distribución de entre pares de calificaciones** (Figura 1). Este gráfico muestra la relación de los estudiantes con respecto a las calificaciones entre parciales. Muestra en distinto color los alumnos por sexo para valorar cualquier relación. Es útil para conocer si un grupo de alumnos ha podido perder la progresión durante el curso.

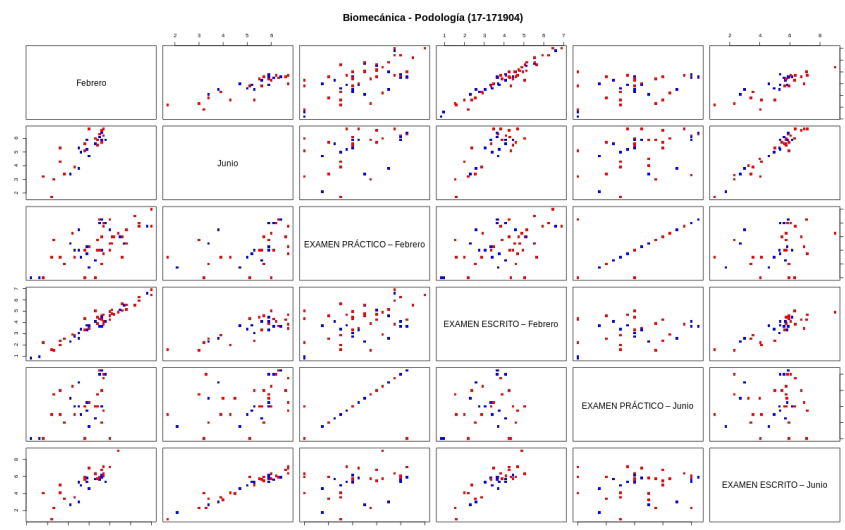


Figura 1. Distribución de entre pares de calificaciones. Ejemplo: Biomecánica, curso 2017/18.

C. **Estudiantes matriculados/aprobados por sexo** (Figura 2). Permite conocer la variación del número de estudiantes matriculados por sexo entre distintos cursos y el tanto por ciento de aprobados por sexo.

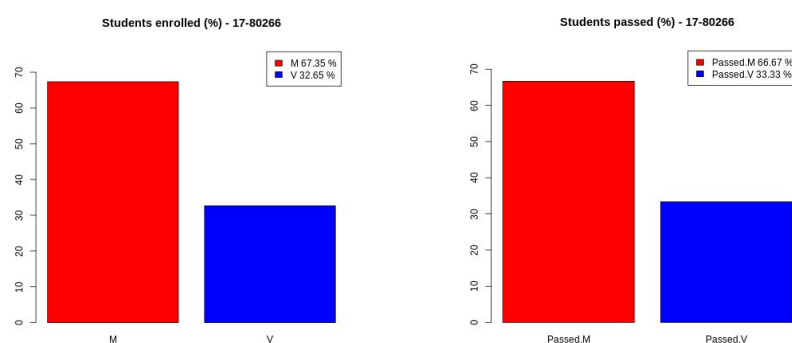


Figura 2. Distribución por sexo de alumnos matriculados y aprobados. Ejemplo: Biología celular, curso 2017/18.

- D. **Distribución de calificaciones por parciales** (Figura 3). Permite comparar la evolución del número de aprobados, suspensos y no presentados por parciales y entre cursos.

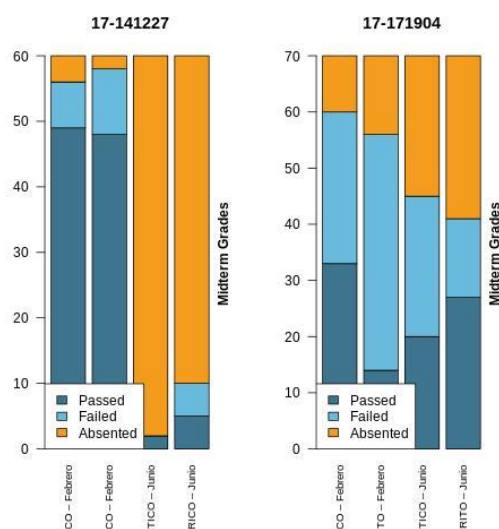


Figura 3. Distribución de calificaciones. Ejemplo: Anatomía III y Biomecánica, curso 2017/18.

- E. **Accesos al curso por fechas** (Figura 4). Para cada curso indicado, se muestra una gráfica de accesos por fecha con referencia a las fechas de los parciales (rojo) y de las autoevaluaciones (gris) si hubieran sido indicadas.

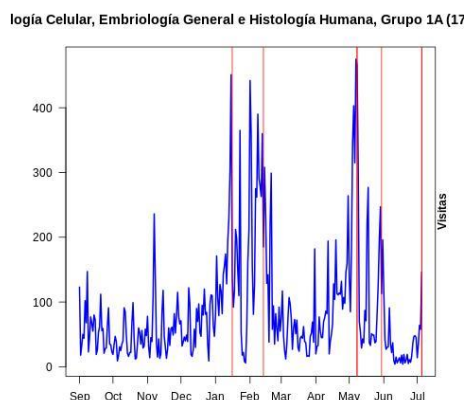


Figura 4. Acceso de alumnos al CV a lo largo del curso. Ejemplo: Biología celular, curso 2017/18.

- F. **Acceso a actividades dentro del curso** (Figura 5). Para cada una de las actividades propuestas (Wikis, Autoevaluaciones, Glosarios, Accesos URL

externas, Foros, Tareas y Recursos). Muestra, por fechas, las actividades realizadas dentro del curso. Al igual que la anterior gráfica, muestra como referencia las fechas de los parciales y las autoevaluaciones y la media de recursos descargados por día durante el periodo estipulado (<fec-ini> y <fec-fin>).

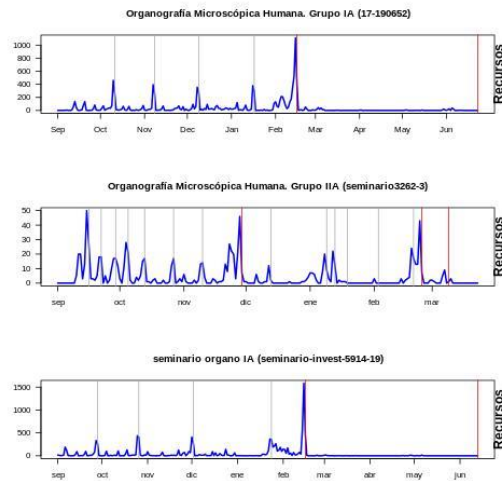


Figura 5. Acceso de alumnos a recursos del CV a lo largo del curso. Ejemplo: tres espacios de prácticas de OMH, cursos 2017/18 y 18/19.

G. **Clúster de estudiantes por nota obtenida** (Figura 6). Segmenta los resultados finales en tres rangos (de 0 a 5, de 5 a 8 y de 8 a 10) mostrando, mediante un diagrama de rádar o estrella, el uso de las herramientas que realiza cada clúster de estudiantes.

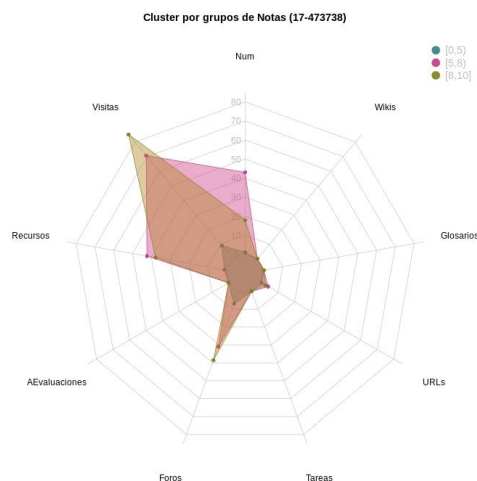


Figura 6. Clústeres de alumnos según nota final. Ejemplo: Salud laboral, curso 2017/18.

H. **Diagrama de cajas por parciales** (Figura 7). Permite observar en una sola gráfica la distribución de las calificaciones por cada parcial (cuartiles, rango intercuartílico y valores atípicos). Con este gráfico podemos identificar alumnos con calificaciones

excepcionales en ambos sentidos y comparar los valores entre los distintos parciales.

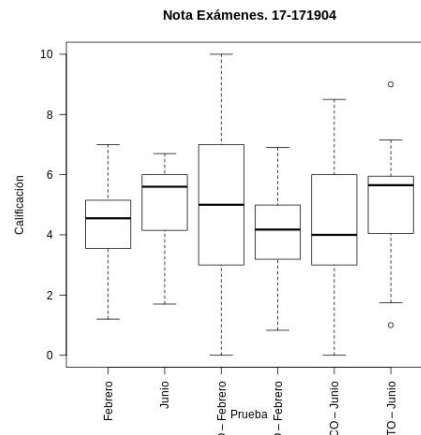


Figura 7. Diagrama de caja de calificaciones. Ejemplo: Biomecánica, curso 2017/18.

- I. **Comparativa de accesos con la media de las asignaturas virtualizadas en el CV (Figura 8).**

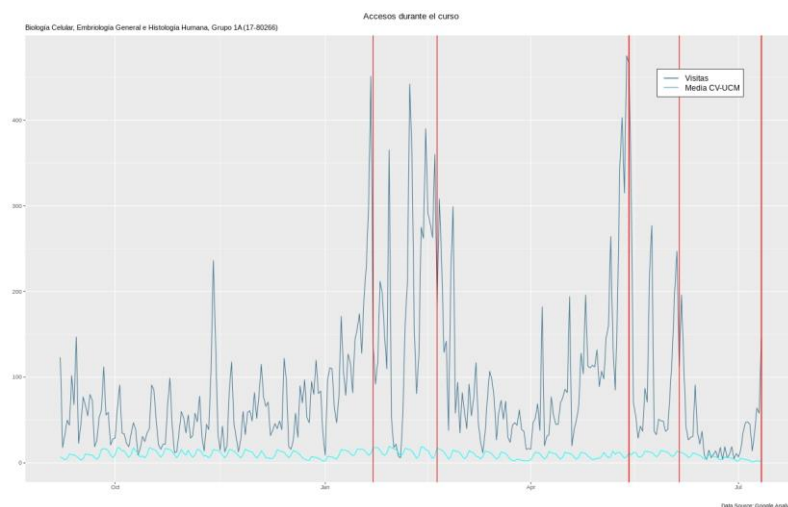


Figura 8. Acceso de alumnos al CV a lo largo del curso comparado con media UCM. Ejemplo: Biología celular, curso 2017/18.

- J. **Accesos por día de la semana (Figura 9).** Este gráfico permite valorar los accesos de los usuarios por día de la semana y compararlo con los accesos del resto de asignaturas dentro del CV de la universidad. Este gráfico, junto con el siguiente, permite identificar patrones de acceso de los alumnos y conocer sus ritmos de trabajo.

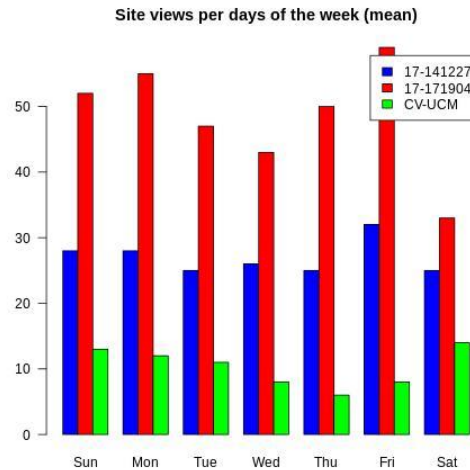


Figura 9. Acceso de alumnos al CV a lo largo de la semana comparado con media UCM. Ejemplo: Anatomía III y Biomecánica, curso 2017/18.

K. **Accesos a la asignatura por franja horaria** (Figura 10). Permite conocer a qué hora se conectan con más frecuencia los estudiantes al CV y programar actividades acordes a esos ritmos.

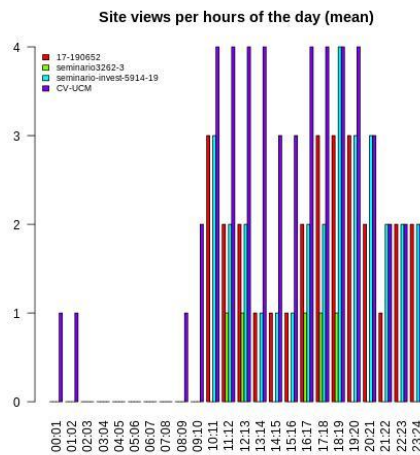


Figura 10. Acceso de alumnos al CV a lo largo de la semana comparado con media UCM. Ejemplo: tres espacios de prácticas de OMH, cursos 2017/18 y 18/19.

L. **Mapa de Calor de la asignatura** (Figura 11). Muestra un gráfico para conocer en qué días de la semana y segmento horario la asignatura tiene más actividad.

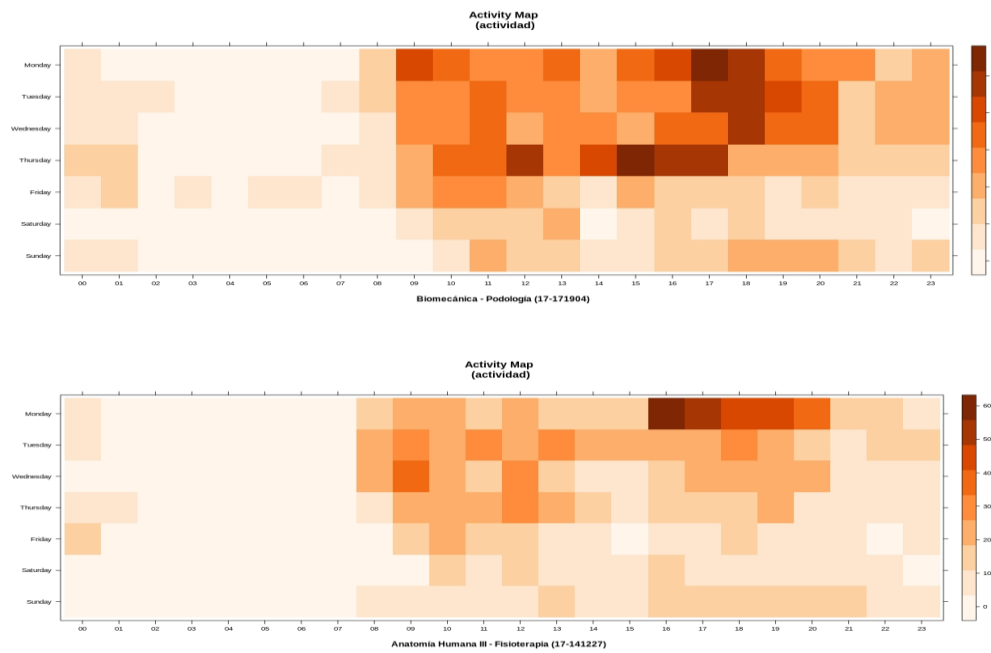


Figura 11. Mapas de actividad en CV. Ejemplo: Anatomía III y Biomecánica, curso 2017/18.

M. **Árboles de decisión** sobre distintas calificaciones, como la nota final (Figura 12). Los árboles de decisión nos ayudan a tomar la decisión “más acertada”, desde un punto de vista probabilístico, ante un abanico de posibles decisiones. En este caso sirve para evaluar cómo afecta probabilísticamente la participación de cada alumno en cada una de las herramientas disponibles en el curso de cara a la nota final obtenida.

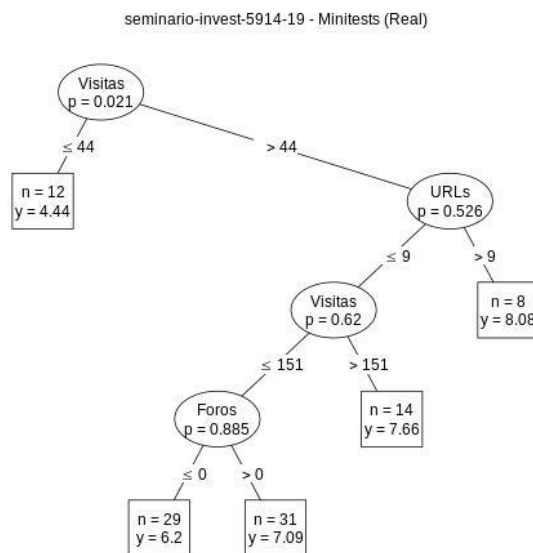


Figura 12. Árbol de decisión. Ejemplo: espacio de prácticas OMH, curso 2018/19.